International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Measuring and Comparing Semantic Structure of Ontology

Nirmitee N. Kurhekar[1], Prof. L. J. Sankpal[2]

[1]Department of Computer Engineering, Sinhgad Academy of Engineering, Kondhwa, Pune, India

Abstract: ***Ontologies which are already constructed can be reused. Ontology reuse offers great benefits by measuring and comparing ontologies. A fundamental operation in ontology reuse is to compute similarity and dissimilarity among ontological entities such that one can establish certain level of correlation between ontological entities used in different ontologies by predefined measures and semantic comparison. Using Graph Derivation Based Approach, semantic structure of two or more ontologies can be compared and measured. In this paper, the novel method called as Graph Walk Based Method is used to compare and measure the semantic structure of the ontologies. This method compares the similarity with polysemy and toponymy. It re-rank the path covered in graph walk. It integrates the textual and structured elements of the data. The graph representation is modular and it is straightforward to extend the graph schema to include other information sources available, such as concepts that represent activities, relations that correspond to domain, etc. While the graph walk and weight tuning effectively model concept associations, re-ranking allows further flexibility in adapting the generated similarity measure to different tasks using high-level and task-specific information.***

Keywords: Ontology, Semantic Structure, Re-ranking, Ontology Measures

1. Introduction

a) **Idea:** Ontology can be defined as the Explicit Specification of conceptualization. In other words, we can define Ontology as an abstract view of set of concepts and their relationships. It gives the semantic structure of any concept related to the specific domain. In todays internet world, information retrieval is one of the important things. For appropriate information retrieval, knowledge representation and knowledge management must be done accurately. For this purpose ontology can be used eficiently. Ontologies are domain specific and give the complete idea about the particular domain correctly.In the concept of information retrieval, user should get accurate domain-specific information relevant to the query given. Hence to represent the domain, ontologies are used. It is the graphical view of domain. It gives the design of concepts which are semantically related to each other. Ontology starts with the generalize concept and proceeds to more specific concept relevant to the users need.

b) **Motivation:** Construction of ontology plays an important role in retrieval process. But ontology construction is very tedious and cumbersome job. To construct an ontology various algorithms can be used, e.g. Graph Derivation Based Approach. To avoid the problem of ontology construction the concept of Ontology Reuse is evolved. This means that the existing ontology of relevant concept can be taken into consideration. Also according to the users need some modifications can be done in that structure. This helps to reduce work and time in generating new ontology which makes the retrieval process faster.

In this paper, method of ontology reuse is proposed. Here, two already existing ontologies are compared for their similarity and dissimilarity. According to the comparison, ontologies are re-ranked using graph-walk method. Generally, ontology incorporates the textual data only. But here images are also taken into consideration while generating the new ontology or comparing two existing ontologies.

2. Related Work

1. A cluster-based measure combines the minimum path length and the taxonomical depth and defines clusters for each of the branches in the hierarchy with respect to the root node. This research explores a new way to measure semantic similarity between biomedical concepts using multiple ontologies. We propose a new ontology-structure-based technique for measuring semantic similarity in single ontology and across multiple ontologies in the biomedical domain within the framework of unified medical language system (UMLS). The proposed measure is based on three features: 1) cross-modified path length between two concepts; 2) a new feature of common specificity of concepts in the ontology; and 3) local granularity of ontology clusters.
2. An ontology-based measure utilizing taxonomical features was proposed without using tuning parameters to weight the contribution of potentially scarce semantic features. It combines the idea of two popular semantic similarity calculation approaches: graph-based approaches and feature-based measures. In the context of computing semantic similarity, adopted a similarity function to determine similar entity classes by a matching process based on synonym sets, semantic neighbourhoods, and distinguishing features.
3. "Fuzzy measures on the gene ontology for gene product similarity" presented several novel measures for computing the similarity of two gene products with graph-based ontology terms annotated by common taxonomy terms. The fuzzy measure similarity (FMS) has the advantage that it takes into consideration the context of both complete sets of annotation terms when computing the similarity between two gene products. When the two gene products are not annotated by common taxonomy terms, we propose a method that avoids a zero similarity result. To account for the variations in the annotation reliability, we propose a

similarity measure based on the Choquet integral. These similarity measures provide extra tools for the biologist in search of functional information for gene products.

4. "Semantic oriented ontology cohesion metrics for ontology based systems" proposed several measures for ontology cohesion measurement by using implicit semantic information. It proposed a set of ontology cohesion metrics which focuses on measuring (possibly inconsistent) ontologies in the context of dynamic and changing Web. They are: Number of Ontology Partitions (NOP), Number of Minimally Inconsistent Subsets (NMIS) and Average Value of Axiom Inconsistencies (AVAI). These ontology metrics are used to measure ontological semantics rather than ontological structure. This paper can be used as a very useful complementarily of existing ontology cohesion metrics.
5. "Representing ontologies using description logics, description graphs, and rules" presented an interesting work that uses the notion of description graph model to address the problem of insufficient expressivity of describing explicit structured objects in ontological knowledge bases. It proposed an extension of DL languages with description graphs - a knowledge modelling construct that can accurately describe objects with parts connected in arbitrary ways. Furthermore, to enable modelling the conditional aspects of structured objects, it extended DLs with rules. It presented an in-depth study of the computational properties of such formalism. In particular, it identified the sources of undesirability of the general, unrestricted formalism.
6. "A Graph Derivation Based Approach for Measuring and Comparing Structural Semantics of Ontologies" includes the GDR based approach, a three-phase process to transform an ontology to its GDR. Also analyze important properties of GDRs based on which stable semantic measurement and comparison can be achieved successfully and compare GDR based approach with existing graph based methods using a dozen real world exemplar ontologies.

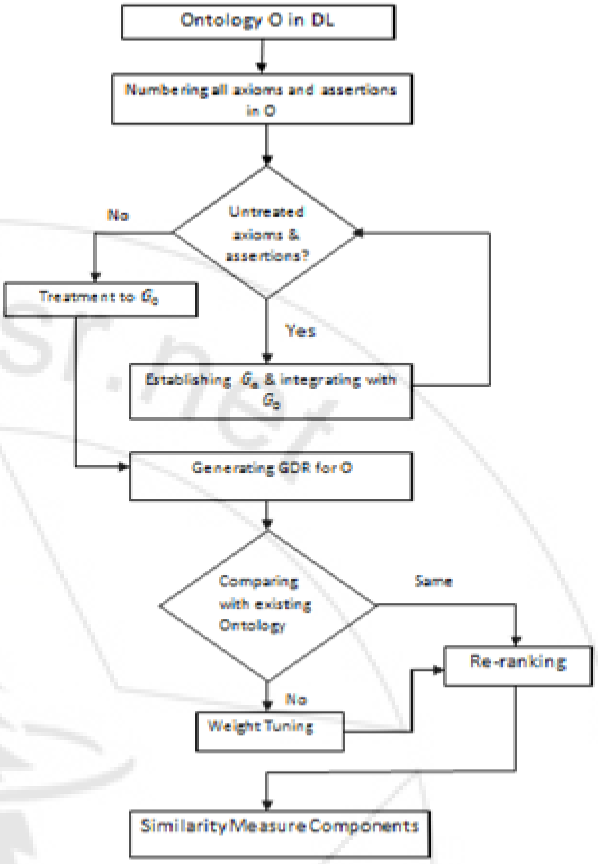
3. Proposed System

3.1 Goal and Objective

Goal of this system is to compare and measure structural similarity of different ontology. By comparing, we are going to find the relevant association of concepts. Objective is to optimize the performance by applying re-ranking to improve the graph walk results and to neglect the problems of polymorphism of ontology representation and the addition of implicit semantic knowledge. As system reuses the existing ontology, it reduces the work and time in generating the new Ontology. It re-ranks the path covered in graph walk. The graph representation is modular and it is straightforward to extend the graph schema to include other information sources available. While the graph walk and weight tuning effectively model concept associations, re-ranking allows further flexibility.

3.2 Methodology

The basic working of the system is depicted as follows:



1. **GDR**: Referring to the axioms & assertions, empty ontology is converted to the GDR which makes the comparison easy & modular.
2. **Graph walk based method**: As the name suggest, using this method one can visit each node of the graph. Also it will find the alias name similarity by comparing two ontologies. After the comparison, initial re-ranking is done using path covered in graph walk.
3. The main advantage of this is it integrates the textual and the structured elements of the data. Previously the ontology comparison is done with the textual data only. But using this technique, two or ontologies with images can easily be compared.
4. For the comparison purpose, firstly parts images are extracted to build the ontology or modifications can be done in the existing ontology. This will be performed by using concept of Image Processing such as calculation of pixels, finding its position, etc….
5. **Re-ranking:** Two ontologies will be compared to check to what extent two ontologies are same. If the compared ontologies are same then re-ranking is done to clarify the result.
6. In this, the feature vectors for the top K nodes retrieved that are to be reranked. This means that feature extraction takes place after the graph walk is completed. Given the set of connecting paths to each of the top K nodes, extracted via the path unfolding procedure, it is straightforward to derive the feature values.
7. If this will not be the case, then weight tuning of the existing ontology is done which will again re-ranked and similarity measures are find out.

4. Conclusion

A GDR derivation based approach is used to stably measure and compare ontologies. By theoretical analysis of the properties of GDR, it is that the GDR of ontology is semantic-preserving and "unique" in terms of labels, connecting structure and isomorphism, which guarantees stable semantic ontology measurement. A general system structure is developed for inducing adaptive similarity measures in divergent data represented as an entity-relation graph. It is developed on graph-walk based paradigms that generate measures of structural similarity between entities in the graph. Here, re-ranking is suggested as a method for learning to rank graph nodes that can use high-level information about the graph walk process. In particular, generic features are proposed that describe the set of paths traversed in reaching a target node from the query of the user.

References

[1] Yung-Shen Lin, Jung-Yi Jiang, and Shie-Jue Lee, "A Similarity Measure for Text Classification and Clustering" IEEE Transactions on knowledge and data engineering, vol. 26, 2014.
[2] Yuhua Li, Zuhair A. Bandar, and David McLean, "An Approach for Measuring Semantic Similarity between Words Using Multiple Information Sources" IEEE Transactions on knowledge and data engineering, vol. 15, 2003.
[3] M. Popescu, J. M. Keller, and J. A. Mitchell, "Fuzzy measures on the gene ontology for gene product similarity," IEEE/ACM Trans. Comput. Bio. and Bioinfo., 2006.
[4] H. Al-Mubaid and H. Nguyen, "Measuring semantic similarity between biomedical concepts within multiple ontologies," IEEE Trans. Syst., Man, Cybern. C, 2009.
[5] D. Sanchez, M. Batet, D. Isern, and A. Valls, "Ontology-based semantic similarity: A new feature-based approach," Expert Syst. Applicat 2012.
[6] A. Rodriguez and M. Egenhofer, "Determining semantic similarity among entity classes from different ontologies," IEEE Trans. Knowl. Data Eng., 2003.
[7] Z. Zou, J. Li, H. Gao, and S. Zhang, "Mining frequent subgraph patterns from uncertain graph data," IEEE Trans. Knowl. Data Eng., 2010.

Author Profile

Nirmitee N. Kurhekar received the B.E. degree in Computer Science & Engineering from H.V.P.M's College of Engineering & Technology, INDIA in 2012 and is pursuing M.E. in Computer Engineering at Sinhgad Academy of Engineering, Pune.

Prof. L. J. Sankpal received the B.E. degree in Computer Engineering from Marathwada University in 1991, M.E. degree in Computer Engineering from Shivaji University in 2001 and persuing Ph D from Bharti University. She is currently working as a Assistant Professor in Sinhgad Academy of Engineering, Pune.